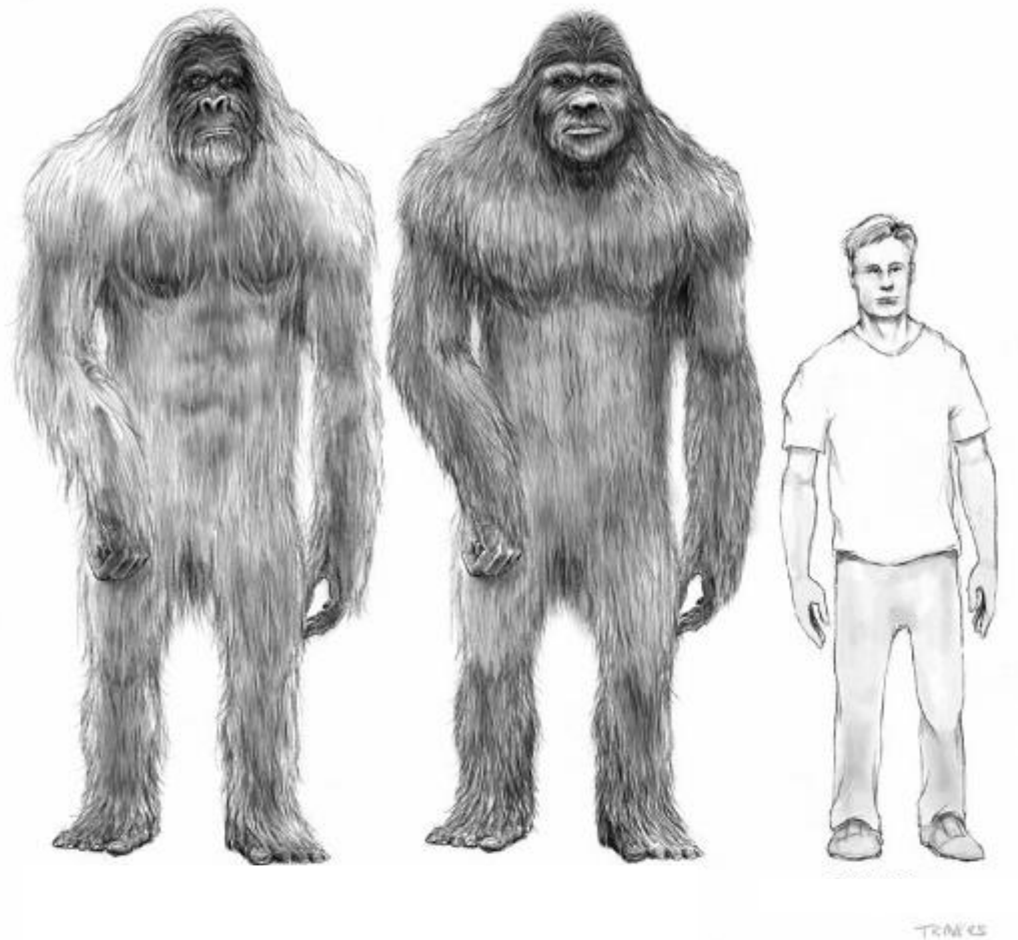
CENTRALE LILLE

# Analyse & traitement de données : Validité et fiabilité

Mise à jour du 7 novembre 2011



Dernière version des diapos disponible ici : Validité et fiabilité

Cours distribué sous licence **Creative Commons,**
selon les conditions suivantes : 


*bachelet@bigfoot.com*
**École Centrale de Lille**
**Villeneuve d’Ascq - France**

Image: source

Source des images indiquées au-dessous ou en cliquant sur l’image

VISIT…

LANZAROTE Caliente.COM

# Comment utiliser ce cours :

1. Mettre les diapos en format plein écran en cliquant sur 
2. Faire défiler l'animation en cliquant **sur** les diapositives

(attention : cliquer sur une image ou un lien ouvre la page web correspondante)



# Tester des hypothèses

Exemple : Le réchauffement climatique est-il une certitude ? Si oui d'où vient-il ?

1. « Il n'y a pas de réchauffement climatique » => réfuté avec de plus en plus de certitude d'année en année (désormais >95%)
2. Autres hypothèses
   1. Le réchauffement climatique est lié à l'activité solaire => réfuté
   2. Le réchauffement climatique est lié à l'activité humaine =>
   3. …

Cf : travaux du GIEC


Utilisation ou copie interdites sans citation

# A la recherche de la vérité …

Pour trouver la vérité, faut-il tester des hypothèses, jusqu'à faire *« sortir la vérité du puits »* ?

Qui ne veut pas voir « la vérité toute nue » ?



La Vérité sortant du puits
Edouard Debat-Ponsan,
musée de l'Hôtel de ville d'Amboise
Domaine Public

# Conjectures et réfutations

On ne prouve pas qu'une théorie est vraie, on prouve que les théories concurrentes sont fausses

⇒ La théorie « vraie » est celle qui

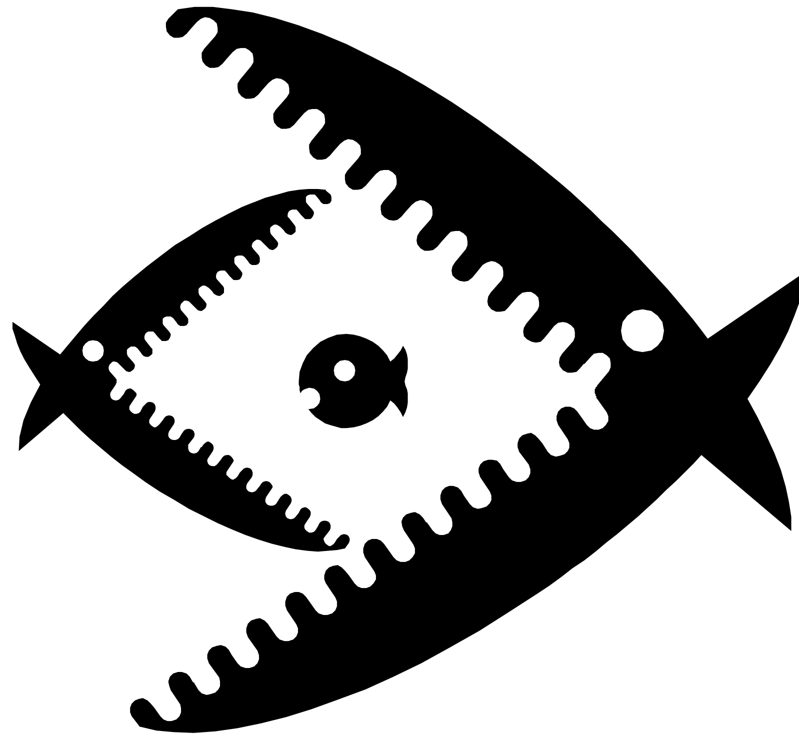
1. Explique un phénomène le mieux possible …
2. permet de formuler des hypothèses réfutables : *principe criticiste*
3. … mais **pas encore** réfutées : *principe faillibiliste*

- Autrement dit, il n'y a pas de théorie vraie, mais seulement une « meilleure théorie » qui est :
  - parmi celles qui sont proposées….
  - … la plus « élégante » (pouvoir explicatif + rasoir d'Occam) …
  - …. dont on dispose à un moment donné, en attendant… (K. Popper)


Utilisation ou copie interdites sans citation

# Il n'y a pas de vérité scientifique, mais une population de théories en concurrence






Utilisation ou copie interdites sans citation

# Validité et Fiabilité

Lorsque l'on mesure une grandeur, on souhaite que les résultats soient utiles et reflètent bien les concepts qu'elle représente.

À cet effet, on a deux préoccupations :

1. Le dispositif de recherche doit être **valide**
2. La mesure doit être **fiable**


Bibliographie / sources :

- William D. Richards http://www.uwm.edu/People/hessling/p325/Psy325_Chap3.pdf
- Bob Hessling http://www.sfu.ca/~richards/Zen/show4/ch4.html
- Igalens et Roussel, « Méthodes de recherche en management », chap 10




Utilisation ou copie interdites sans citation

# La Validité

Les trois exigences de validité

1. Validité interne
2. Validité externe
3. Validité de construit / concept

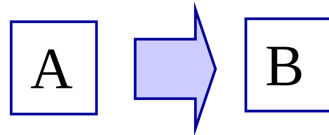


Utilisation ou copie interdites sans citation

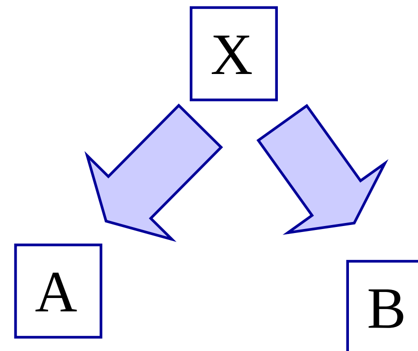
# Validité interne

Validité interne = On a bien identifié cause et effet

Autrement dit, pas d'explication rivale : les variations de la variable à expliquer sont causées uniquement par les variables explicatives




? ou





Utilisation ou copie interdites sans citation

# Qu'est-ce que la causalité ?

Selon John Stuart Mill (1806-1873), trois critères permettent d'inférer la causalité :

i. La covariation,
   - Cause et effet sont corrélés
ii. La précédence temporelle
   - La cause précède l'effet
iii. L'élimination d'explications alternatives.
   - Pas de troisième variable



*Trouvez huit biais limitant cette validité interne … et les moyens de les éviter (application à un questionnaire d'intention d'achat d'une population dans un centre commercial...)*

John Stuart Mill, 1865
John Watkins (1823-1874)
Domaine Public

# Huit biais limitant la validité interne (1/2)

| Type | Origine | Comment l'éviter |
|---|---|---|
| **Effet d'histoire** | Des évènements extérieurs à l'étude faussent les résultats (Noël) | • Examen critique de la période d'étude<br>• Réduire la période d'étude |
| **Effet de maturation** | Les individus ont changé pendant l'étude (réussite/échec au bac, entre ou sort du centre commercial) | • Examen critique des individus<br>• Réduire la période d'étude |
| **Effet de test (not. pour une étude longitudinale/ par panel)** | Les réponses au deuxième questionnaire sont affectées par le fait d'avoir déjà répondu (mémoire) | • Ne pas questionner deux fois les mêmes individus *(?)*<br>• Éviter la mémorisation ? |
| **Effet d'instrumentation** | Les questions utilisées pour recueillir les données sont mal formulées (mots compliqués….) | • faire valider le questionnaire par un expert<br>• protocole normalisé<br>• Méthode de la pensée à voix haute.. |

Rémi BACHELET

Campbell DT, Stanley JC (1966). “Experimental and quasi-experimental designs for research”.

# Huit biais limitant la validité interne (2/2)

| Type | Origine | Comment l'éviter |
|---|---|---|
| **Effet de régression statistique** | Présélection des individus sur la base de caractère extrêmes (les participants avec les plus mauvais résultats à un test sont inclus dans un programme de formation, ils réussissent au prochain test peu importe le programme, … parce que les probabilités d'obtenir d'aussi mauvais résultats deux fois sont faibles) | • Revoir la constitution de l'échantillon |
| **Effet de sélection** | L'échantillon n'est pas représentatif de la population pertinente (les enquêteurs sont actifs la journée, mais plus après 20h… les PCS faisant les courses tard sont oubliées) | • Attention au plan de collecte ! |
| **Effet de mortalité expérimentale (pour une étude longitudinale/ par panel)** | Des sujets disparaissent en cours d'étude (des participants abandonnent l'étude) | • Remplacer les sujets perdus<br>• Optimiser les moyens de garder le contact. |
| **Effet de contamination** | Un individu interrogé apprend à l'avance par les autres l'objet de l'étude ou les réponses attendues | • Cacher objectif/les moyens de l'étude<br>• Mener l'étude rapidement éviter la communication entre sujets |

Rémi BACHELET

Campbell DT, Stanley JC (1966). "Experimental and quasi-experimental designs for research".

# Validité externe

Validité externe = Les résultats sont-ils généralisables ?

Autres situations : pays, époques, populations ….

Un risque :

- « Overfit » : en recherche inductive, on construit un modèle à partir de données… et parfois celui-ci *colle* tellement à ces données … qu'il ne représente plus le phénomène qu'on voulait étudier au départ…. mais uniquement les données.

Une faible validité externe signifie que l'on ne peut pas généraliser, *mais la recherche peut tout de même être intéressante* (on a très bien étudié **un** terrain donné).

Toujours revenir sur la validité externe en conclusion d'une étude (« limites et perspectives »).


Utilisation ou copie interdites sans citation

# Validité interne et externe s'opposent-elles ?

À votre avis ?


Utilisation ou copie interdites sans citation

# Validité interne et externe s'opposent-elles ?

À votre avis ?

- la validité interne demande un meilleur contrôle des situations observées
- la validité externe demande des situations, méthodes et terrains de collecte plus variés et ouverts

➢ Plus on multiplie les situations, plus le contrôle qu'on peut exercer sur les biais est faible.

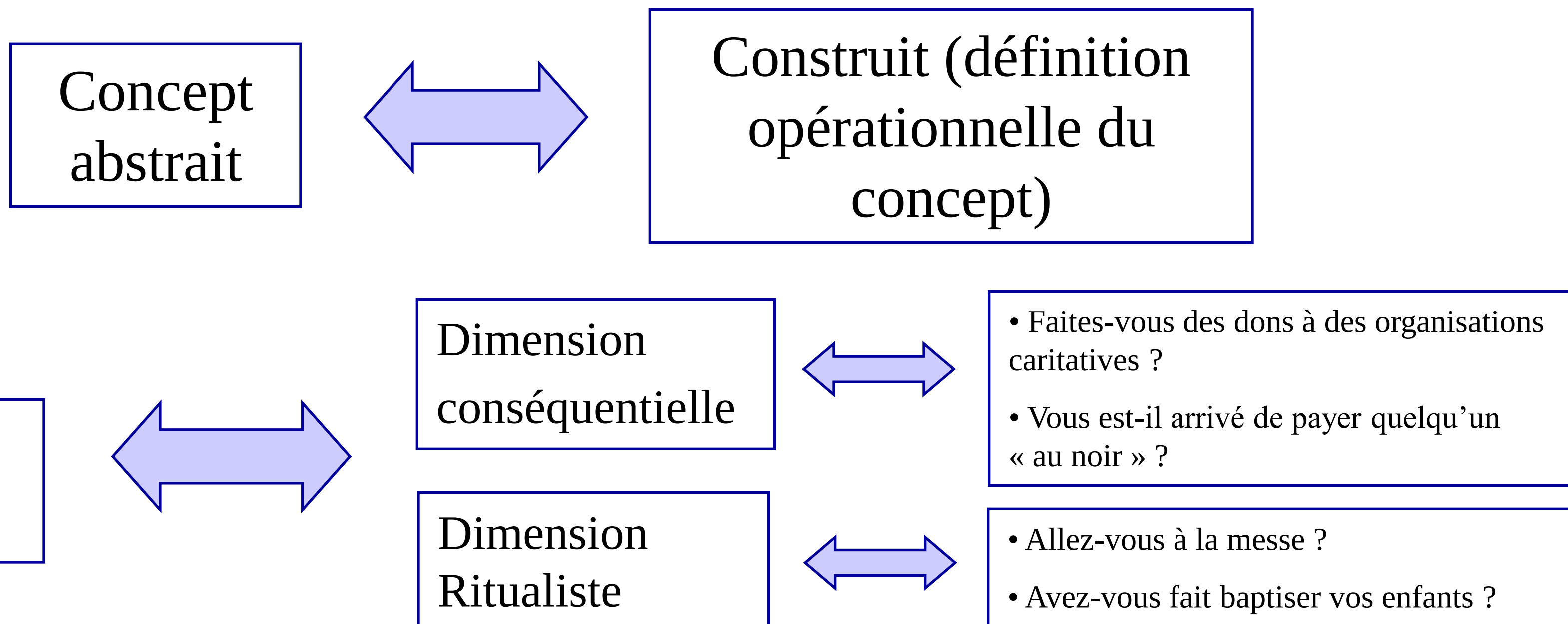
Donc … Oui


Utilisation ou copie interdites sans citation

# Validité de construit / concept

Validité de construit = on mesure bien ce que l'on veut mesurer

- Les recherches en sciences humaines portent sur des concepts abstraits non toujours directement observables (*le changement, l'autonomie, les savoirs-être…* ).
- Les théories étudiées reposent donc sur des concepts, souvent polysémiques et difficiles à définir.
- Il faut donc valider la relation :





Utilisation ou copie interdites sans citation

# Les critères de validité de concept / construit

On compte jusqu'à douze moyens de l'évaluer, retenons notamment

- La validité de contenu
  L'opérationnalisation représente le concept sous tous ses aspects
- La validité d'observation
  Degré auquel le concept étudié peut se réduire à des observations
- La validité de critère
  Degré auquel on peut affirmer que le construit opérationnalisé est corrélé au concept qu'il est sensé représenter.
- La validité de trait
  Degré auquel on peut affirmer que le construit opérationnalisé permet de mesurer le concept qu'il est sensé représenter.
- La validité discriminante
  Renvoie à la sensibilité/spécificité de l'échelle de mesure.
- La **validité convergente**
  Deux mesures du concept par deux méthodes différentes sont efficaces.
- La **validité divergente**
  Degré auquel le concept diffère d'autres concepts proches
- La validité systémique
  Degré auquel le concept permet l'intégration de concepts antérieurs ou de produire de nouveaux concepts
- …

Voyons comment mettre en évidence les validités convergente et discriminante



Zaltman, Pinson, and Angelmar, 1973 "Metatheory and Consumer Research" amendement de S de Chalvron pour la validité discriminante

# Évaluer la validité de construit : Matrice multitraits-multiméthodes

La matrice multitraits-multiméthodes synthétise la corrélation entre les mesures de différents traits par des méthodes différentes

| | | Method 1 – Experiment 1 | | Method 2 – Experiment 1 | | Method 3 – Experiment 1 | |
|---|---|---|---|---|---|---|---|
| | | U1 | A1 | U2 | A2 | U3 | A3 |
| **Method 1 – Experiment 1** | | | | | | | |
| Apparent Usability | $U_1$ | ( ) | | | | | |
| Aesthetics | $A_1$ | .92 | ( ) | | | | |
| **Method 2 – Experiment 1** | | | | | | | |
| Apparent Usability | $U_2$ | .92 | .79 | ( ) | | | |
| Aesthetics | $A_2$ | .87 | .90 | .83 | ( ) | | |
| **Method 3 – Experiment 1** | | | | | | | |
| Apparent Usability | $U_3$ | .84 | .77 | .90 | .89 | ( ) | |
| Aesthetics | $A_3$ | .81 | .85 | .79 | .94 | .92 | ( ) |

**=1** sur la diagonale

On peut aussi corréler les traits et méthodes avec eux-mêmes (plusieurs expériences). On évalue ainsi la reproductibilité d'une méthode …. **= sa fiabilité. cf page suivante**

multitraits-multiméthodes
**= validité discriminante**

diagonale monotrait-multiméthodes
**= validité convergente**

Campbell DT, Stanley JC (1966). "Convergent and discriminant validation by the multitrait-multimethod matrix".

Image d'origine: http://www.acm.org/sigchi/chi97/proceedings/paper/nt-tb3.gif

# La Fiabilité *(reliability)*

La mesure est-elle régulière ?

La question de la fiabilité de l'instrument de mesure



Utilisation ou copie interdites sans citation

# Qu'est-ce que la fiabilité?

Les opérations de la recherche peuvent être répétées à un autre moment par un autre chercheur avec le même résultat.

Une mesure parfaite donne le même résultat à chaque fois qu'elle est refaite.

- Trop de variabilité (écart-type) entre mesures = instrument pas fiable


Utilisation ou copie interdites sans citation

# Mesures de fiabilité en recherche quantitative

Mesure lue

=

\+ +




Méthodes d'estimation de la fiabilité

1. Re-retest

   Refaire exactement la même mesure à des moments différents.

2. Méthodes des formes alternatives

   Re-retest avec un test différent (alternatif) mais mesurant la même chose.

3. Mesure de la cohérence interne

   Pour une échelle construite à partir de plusieurs items, tester leur cohérence. Alpha de Cronbach, rho de Jöreskog …


Utilisation ou copie interdites sans citation

# Validité et Fiabilité sont ils réservés à la recherche quantitative ?

Non !

Mais on ne procède pas de la même manière deux ces deux cas

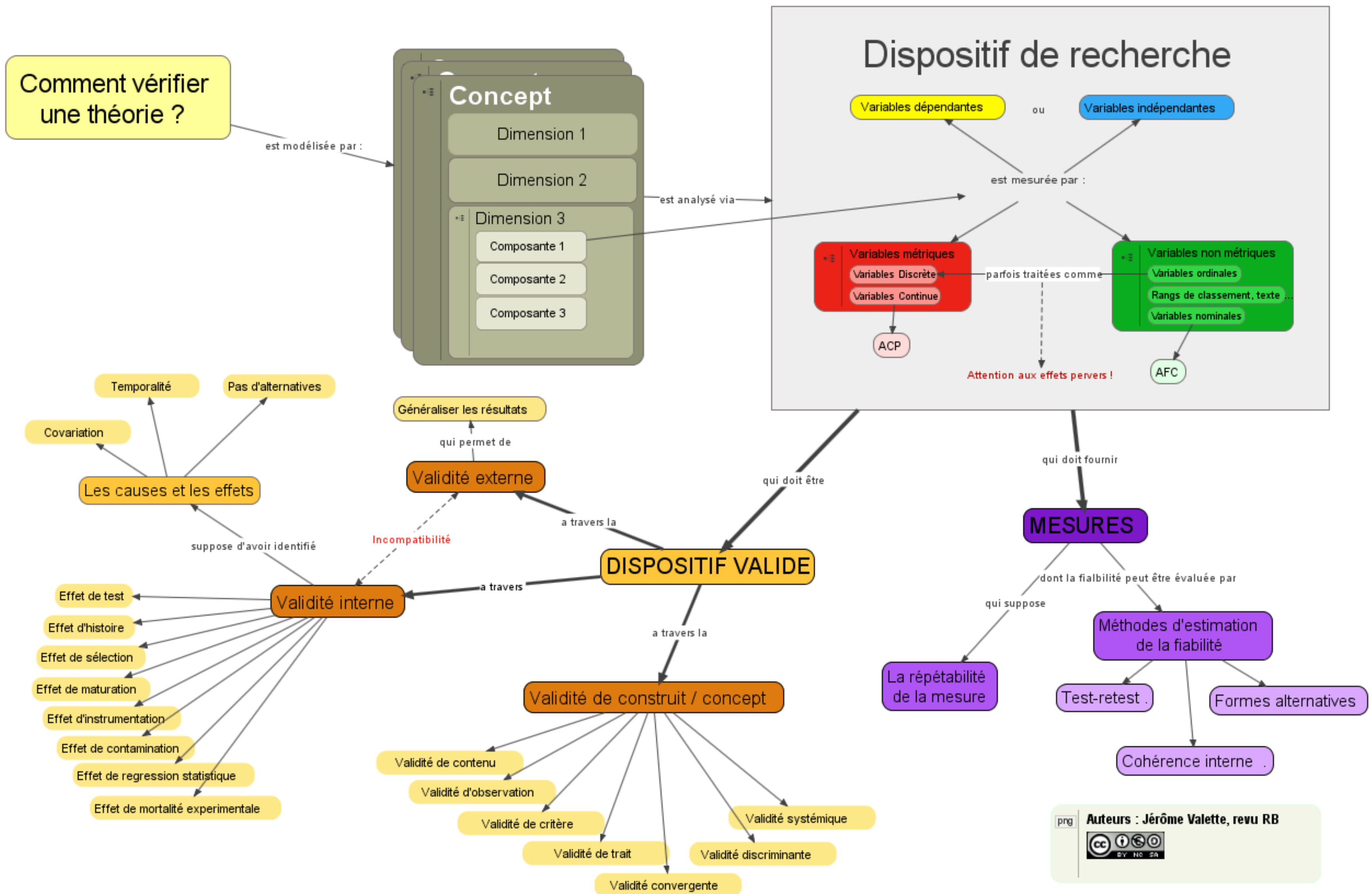
- Quanti => tests statistiques
- Quali => prendre certaines précautions, par exemple vérifier si on conserve le même résultat en changeant d'observateur

Problèmes possibles pour le qualitatif :

- Critères observés ambigus
- Mal former les observateurs/enquêteurs problème d'harmonisation ses données collectées
- Situation trop dure à étudier


Utilisation ou copie interdites sans citation

# Pour mémoriser ce cours, voici une carte conceptuelle, vous trouverez ici d'autres cartes heuristiques sur ce cours






Utilisation ou copie interdites sans citation

# Questions ?



Utilisation ou copie interdites sans citation

# Autres cours :

1. Explorer ou vérifier ? Deux catégories d'approches
2. Éventails des démarches de recueil de données
3. Conception de questionnaires
4. Techniques d'entretien et reformulation
5. L'Analyse Factorielle des Correspondances pour les nuls
6. Validité et Fiabilité des données


Utilisation ou copie interdites sans citation



# Rémi BACHELET

*Enseignant-chercheur,*
*Ecole Centrale de Lille*

*Mon CV est disponible ici.*

**Mes principaux cours à Centrale**
Gestion de projet, sociologie des organisations, recueil, analyse et traitement de données, prévention du plagiat, module de marchés financiers, cours de qualité et méthodes de résolution de problèmes, établir des cartes conceptuelles, utiliser Wikipédia et CentraleWiki, formation au coaching pédagogique et à l'encadrement


Utilisation ou copie interdites sans citation